Paragraph on page 6, at lines 2-7:

B1

Typically (i) is any allelic variant of $P450_{cam}$ of Pseudomonas putida (e.g. of the polypeptide sequence shown in SEQ ID No. 2). Typically (ii) is a species homologue of (i) which has sequence homology with (i), and is typically $P450_{BM-3}$ of *Bacillus megaterium* (e.g. the polypeptide sequence shown in SEQ ID No. 4 and the nucelotide sequence is shown in SEQ ID NO: 3), $P450_{terp}$ of *Pseudomonas sp,* $P450_{eryF}$ of *Saccharopollyspora erythraea,* or P450 105 D1 (CYP105) of *Streptomyces griseus* strains.

Paragraph on page 6, at lines 22-32:

B2

The discussion below provides examples of the positions at which substitutions may be made in $P450_{cam}$. The same substitutions may be made at equivalent positions in the homologues. Standard nomenclature is used to denote the mutations. The letter of the amino acid present in the natural form is followed by the position, followed by the amino acid in the mutant. To denote multiple mutations in the same protein each mutation is listed separated by hyphens. The mutations discussed below using this nomenclature specify the natural amino acid in $P450_{cam}$, but it is to be understood that the mutation could be made to a homologue which has a different amino acid at the equivalent position. Note that the amino acid numbering shown in SEQ ID No. 2 for $P450_{cam}$ does not correspond to the numbering used in the description to denote mutations. The numbering in SEQ ID No. 2 is one more than the numbering in the description for a particular position.

Paragraph on page 9, at lines 4-10:

B3

Alternatively a linker may be present between the components. The linker generally comprises amino acids that do not have bulky side chains and therefore do not obstruct the folding of the protein subunits. Preferably the amino acids in the linker are uncharged. Preferred amino acids in the linker are glycine, serine, alanine or threonine. In one embodiment the linker comprises the sequence N-Thr-Asp-Gly-Gly-Ser-Ser-Ser-C (SEQ ID NO:6). The linker is typically from at least 5 amino acids long, such as at least 10, 30 or 50 or more amino acids long.

Paragraph on page 11, at lines 2-6:

B4

In one embodiment the three genes encoding the three proteins of the $P450_{cam}$ system, i.e. camA, camB, and camC are placed in the mobile regions of standard transposon vectors and incorporated into the genome of Pseudomonas and flavobacteria. Alternatively plasmid vectors for expressing these genes may used, in which case the $P450_{cam}$ gene cluster will be extra-chromosomal.

Paragraph on page 14, at lines 2-18:

Oligonucleotide-directed site-specific mutagenesis was carried out by the Kunkel method (Kunkel, T. A. *Proc. Natl. Acad. Sci. USA 1985*, **82**, 488-492) using the Bio-Rad Mutagen kit. The recommended procedure is summarised as follows. An M13 mp19 subclone of the *camC* gene encoding $P450_{cam}$ (SEQ ID NO: 1) was propagated in the *E. coli* strain CJ236. This strain has the *dut*⁻ *ung*⁻ phenotype and thus will tolerate the inclusion of uracil in place of thymine in DNA molecules. After three cycles of infection, uracil-containing single stranded (USS) M13 DNA was readily isolated by phenol extraction of mature M13 phage particles excreted into the growth medium. The mutagenic oligonucloetide (or oligonucleotides) were phosphorylated with T4 polynucleotide kinase and then annealed to the USS template. The four nucleotides, DNA polymerase, DNA ligase, ATP and other chemical components were added and the second strand was synthesised *in vitro*. The double stranded form thus obtained was transformed into the *dut*+ *ung*+ *E. coli* strain MV1190, which should degrade the uracil-containing template strand and propagate the mutant strand synthesised *in vitro*. Plaques were picked and phages of possible mutants grown in *E. coli* strains MV1190 or TG1. The single-stranded DNA from these were sequenced to determine whether the mutagenesis reaction was successful. The mutagenic efficiency was 50 - 80%.

B5

Paragraph on page 16, at lines 19-22:

For the *camA* gene the primer below (SEQ ID NO: 5) was used at the 5' end of the gene to introduce the *Eco* RI cloning site and to change the first codon from GTG to the strong start codon ATG.

B6

5'- GAG ATT AA**G AAT TC**A TAA ACA CAT G**GG AG**T GCG TGC CAT **ATG** AAC GCA AAC

Eco RI RBS |→*camA*

Paragraph bridging page 16, line 24 to page 17, line 3:

At the 3' end of *camA* the primer (nucleotide sequence is SEQ ID NO: 7) was designed such that 15 bases are complementary to nucleotide sequence of the last five amino acid residues of *camA*. The stop codon immediately after the GCC codon for the last amino acid was removed, and then part of a seven amino acid linker (Thr Asp Gly Gly Ser Ser Ser; SEQ ID NO: 6) which contained a *Bam* HI cloning site (GGATCC = Gly Ser) was introduced. The coding sequence was thus (amino acid sequence is SEQ ID NO:8):

B7

5'- GAA CTG AGT AGT GCC ACT GAC GGA **GGA TCC** TCA TCG-3'

camA → Thr Asp Gly Gly Ser

|*Bam* HI|

Paragraph on page 17, at lines 5-6:

B8

The primer sequence shown below (SEQ ID NO: 9) is the reverse complement used for PCR:
5'- CGA TGA **GGA TCC** TCC GTC AGT GGC ACT ACT CAG TTC-3'

Paragraph on page 17, at lines 8-13:

B9

For the camB gene the primer at the 5' end (nucleotide sequence is SEQ ID NO: 10; amino acid sequence is SEQ ID NO: 11) incorporated the second half of the peptide linker between the reductase and redoxin proteins, and the restriction site *Bam* HI for joining the two amplified genes together.

5'- TCA TCG **GGA TCC** TCA TCG **ATG** TCT AAA GTA GTG TAT-3'
Gly Ser Ser Ser |→ *camB*
|*Bam* HI| Start

Paragraph on page 17, at lines 14-19:

B10

At the 3' end of *camB* the primer incorporates 12 nucleotides complementary to the end of *camB* followed by the stop codon TAA, a 6 nucleotide spacer before the GGAG ribosome binding site. *Xba* I and *Hind* III sites were then added to allow cloning of the *camC* gene when required. The sequence of the coding strand was therefore (SEQ ID NO: 12):

5'- CCC GAT AGG CAA TGG **TAA** TCA TCG **GGAG TCT AGA** GCA TCG **AAG CTT** TCA TCG-3'
CamB →|stop RBS *Xba* I *Hind* III

Paragraph on page 17, at lines 20-21:

B11

The primer shown below (SEQ ID NO: 13) is the reverse complement used for PCR:

5'-CGA TGA AAG CTT CGA TGC TCT AGA CTCC CGA TGA TTA CCA TTG CCT ATC GGG -3'

Paragraph on page 19, at lines 2-6:

B12

We used the internal and unique restriction site *Mlu* I (recognition sequence ACGCGT) within the *camB* gene as the starting point so that the PCR product has a different size from the PCR template fragment. The primers were as follows:

5'- TCA TCG **ACG CGT** CGC GAA CTG CTG-3' (SEQ ID NO: 14)

where the *Mlu* I site is in bold.

Paragraph on page 19, at lines 7-10:

B 13

The desired coding sequence at the 3' end of the *camB* gene (SEQ ID NO:15) was:

5'- CCC GAT AGG CAA TGG **TAA** GTA GGT GAA TAT CTA ATC CCC ATC
camB →|stop

TAT GCG CGA GT**G GAG TCT AGA** GTT CGA-3'
RBS *Xba* I

Paragraph bridging page 20, lines 25 to page 21, line 7:

The primers shown below maintain the *Hind* III cloning site AAGCTT:

5'-TCA TCG **AAG CTT** GGC TGT TTT-3' (SEQ ID NO: 16)
Hind III |→ vector

B 14

At the other end the coding sequence desired was (SEQ ID NO: 17):

5'-ACA ATT TCA CAC **AGGA TCT AGA** C **CAT ATG** TCA TCG **AAG CTT** TCA TCG-3'
Vector →|RBS *Xba* I *Nde* I *Hind* III

This sequence maintained the *Nde* I and *Hind* III sites but the new *Xba* I site was introduced upstream of the *Nde* I site. The PCR primer used was the reverse complement of the desired sequence (SEQ ID NO: 18):

5'-CGA TGA AAG CTT CGA TGA CAT ATG GTC T AGA TCCT GTG TGA AAT TGT-3'